

R58-45

Response Set in Measurement of Food Preference¹

Howard G. Schutz² and Joe Kamenetzky

Quartermaster Food and Container Institute, Chicago

In many surveys of consumer preference, respondents are presented with the names of products and are asked to indicate their degree of liking or dislike for each. During the years of 1950-1954, surveys of over 30,000 military personnel were conducted by this Institute in order to determine the relative preferences for more than 400 commonly served foods. Each respondent rated up to 54 food names on a nine-interval scale (1, 2), the intervals being successively anchored with the following descriptive categories: like extremely, like very much, like moderately, like slightly, neither like nor dislike, dislike slightly, dislike moderately, dislike very much, and dislike extremely. The scale categories were assigned successive integers from 1 to 9 beginning at the *dislike* end, and the ratings then treated quantitatively. Mean ratings fulfilled the major purposes of allowing the items to be rank ordered and of guiding the menu planners in the selection of foods to be served in future meals.

However, several questions arose in the interpretation of the survey results. First, ingredients used in, and methods of preparing most foods are not constant; variability of these factors should be reflected in the variability of the mean preference ratings of different preparations of these items when these are actually served. Over time, most individuals experience different qualities of servings of the same foods, and it is not known whether survey respondents evaluate the foods in terms of some ideal or idealized experience with them, their least favorable experiences, or some "average" experiences. The practical

implications of this problem will be discussed later.

Second, mean taste-test ratings of food items prepared and evaluated under laboratory conditions are typically lower than the corresponding survey means. This fact suggests the hypothesis that survey respondents evaluate foods as they remember the better servings of them.

The primary purpose of the present study was to determine the set or frame of reference survey respondents use in rating food names according to preference.

Method

The respondents were 305 male enlisted personnel attending service schools at the Great Lakes Naval Training Center, Great Lakes, Illinois. As they left the mess hall following the noon meal the men were divided into three groups of approximately 100 men in each. Each group was administered a questionnaire consisting of 54 foods selected from previous menus served at this installation as representing eight food classes: main dishes, desserts, vegetables, soups, beverages, potatoes and starches, breakfast foods, and breads. The foods were listed in random order. All groups rated each food on the scale described above. The groups were randomly given one of three sets of instructions.

Respondents in the first group ($N = 101$) rated the foods under the usual survey instructions which read, in part: "For each food listed in the following pages, circle the reply which tells how much you like or dislike that food." On the final page each was asked to indicate the quality of food servings he thought of when he rated the foods: best, better than average, average, poorer than average, poorest.

Members of the second group ($N = 100$) were asked to rate the "Best Serving" of each of these same foods that they had ever eaten, and those in the third group ($N = 104$) rated

¹This paper reports research undertaken at the Quartermaster Food and Container Institute for the Armed Forces, and has been assigned Number 781 in the series of papers approved for publication. The views or conclusions contained in this report are those of the authors. They are not to be construed as necessarily reflecting the views or indorsement of the Department of Defense.

²Now at Battelle Memorial Institute, Columbus, Ohio.

the "Poorest Servings." In order to insure maintenance of the "Best Serving" and "Poorest Serving" sets, "Rate the *BEST* (or *POOREST*) *SERVING* of each food," was printed at the top and bottom of each of the six pages of the questionnaire. All respondents were instructed to circle a "Not Tried" category for a food if they thought they had never eaten that food.

It was felt that this method is preferable to having each respondent rate the foods under all three instructions. The latter procedure would likely have had the effect of subtly suggesting to the respondents that the rating made under the usual survey instructions should lie between the ratings made under the other two instructions. However, in other instances where a specific instruction is not the standard being investigated, the exaggeration of differences between forms of instructions may not be serious and may even be preferable.

Results and Discussion

The detailed results for each of the 54 foods are too lengthy to be presented here.³ Instead, a summary of the findings is given. First, while 29 individual preference means are higher in the "Best Serving" group than in the normal or survey instruction group, 24 differences were in the opposite direction. On the other hand, with but one exception and one tie in means, the means for the normal instruction group are higher than the "Poorest Serving" means; and with three exceptions, the means for the "Best Serving" group are higher than the "Poorest Serving" means. Thus, there appears to be no difference between the normal instructions and "Best Serving" instructions whereas each differ from the "Poorest Serving" set.

Inspection of Table 1 reveals that these results also obtain when we consider the means of the food groups and the overall means, even though those evaluating the items under the normal instructions stated to the final question that they had rated in terms of

Table 1
Mean Ratings of Food Groups as
Function of Instructions

Food Group	No. of Foods in Group	Instructions		
		Normal	"Best Servings"	"Poorest Servings"
Main dishes	12	6.78	6.76	6.05
Vegetables	10	5.50	5.39	4.90
Desserts	9	7.31	7.42	6.71
Potatoes & Starches	6	6.94	6.97	6.11
Soups	5	5.85	5.74	5.43
Breads	5	7.40	7.43	6.84
Breakfast Foods	4	6.75	6.78	6.03
Beverages	3	7.24	7.13	6.80
All Foods	54	6.58	6.57	5.93

slightly "better than average" servings. It also appears that the percentages of respondents endorsing the "Not Tried" category for any item did not vary as a function of instructions.

If respondents in surveys do rate in terms of what they consider to be "Best Servings," then any departure from this optimum should, of course, result in lower ratings. This conclusion would help explain the previously mentioned fact that ratings of foods evaluated in taste tests are almost always rated lower than when evaluated during surveys since in the latter case the foods are idealized.

The method and data have uses other than serving as aids in interpreting survey ratings. Thus, by inspecting the differences between "Best Servings" and "Poorest Servings," we can select for further investigation those items for which variations of ingredients or methods of preparation are of importance in the determination of preference. Some foods at all levels of preference exhibit large variation as a function of instructions. The relatively large differences between "Best Servings" and "Poorest Servings" for such items having different levels of preference as roast chicken, oven-browned potatoes, and Harvard beets, as well as for such food groups as potatoes, suggest that these might prove to be more acceptable if certain procurement and preparation procedures are followed. In contrast, the extremely small differences between instructions for fresh milk may be due to the high level of quality control practiced in the processing of this product.

This general approach of obtaining ratings under three instructions might be of practical

³ A three page table giving the mean rating, standard deviation, and percentage "Not Tried" for each food under each type of instruction has been deposited with the American Documentation Institute. Order Document No. 3547, remitting \$1.25 for 35-mm. microfilm, or \$1.25 for 6 by 8 in. photocopies.

value in other situations involving judgments in the form of ratings. One example is the case where a certain brand of a product (e.g., a prepared cake mix) is used by consumers under varying conditions. Detection of large differences between instructions as a function of usage would demonstrate the lack of versatility of the product and would point to the necessity of instituting such corrective measures as modifications of the product itself or of instructional programs.

A second case concerns instances where persons are the objects of assessment. A rater's over-all opinion of someone in and of itself gives no estimate of the ratee's variability, either in terms of general performance or on any trait such as "initiative." It is possible that much of the inter-rater differences might be attributed to the fact that some evaluate an individual at his best, some at his worst, and others somewhere between these extremes. Having the raters perform multiple ratings might enable: (a) adjusting the raters' evaluations according to their frames of reference; (b) detecting those ratees whose performances vary markedly.

Summary

Over 300 military personnel, assigned randomly to each of three groups were asked to indicate their degree of liking for 54 food items belonging to eight food types. Members of the first group rated each item under customary instructions, those in the second rated the "Best Servings" they ever ate of these same foods, and those in the third rated the "Poorest Servings." The results suggest that regardless of food type, food items are evaluated in terms of the most favorably remembered experiences with them. Some practical implications of the approach used in this study are discussed.

Received September 27, 1957.

References

1. Jones, L. V., Peryam, D. R., & Thurstone, L. L. Development of a scale for measuring soldiers' food preferences. *Food Res.*, 1955, 20, 512-520.
2. Peryam, D. R., & Pilgrim, F. J. Hedonic scale method of measuring food preferences. *Food Technol.*, 1957, 11 (9), Supplement, 9-14.